



Second-order differentiability of probability functions

Wim van Ackooij, Jérôme Malick

► To cite this version:

Wim van Ackooij, Jérôme Malick. Second-order differentiability of probability functions. Optimization Letters, 2017, 11 (1), pp.179-194. 10.1007/s11590-016-1015-7. hal-01486517

HAL Id: hal-01486517

<https://hal.science/hal-01486517>

Submitted on 9 Mar 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Second-order differentiability of probability functions

Wim van Ackooij · Jérôme Malick

Received: date / Accepted: date

Abstract In this paper, we study second-order differentiability properties of probability functions. We present conditions under which probability functions involving nonlinear systems and Gaussian (or Student) multi-variate random vectors are twice continuously differentiable. We provide an expression for their Hessian that can be useful in numerical methods for solving probabilistic constrained optimization problems.

Keywords Stochastic optimization · probabilistic constraints · joint-chance-constraints · differentiability of probabilistic function · nonlinear optimization

Mathematics Subject Classification (2000) 49M37, 65K05, 90C15

1 Introduction

Given a mapping $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^k$ and a random vector $\xi \in \mathbb{R}^m$, a probability function $\varphi : \mathbb{R}^n \rightarrow [0, 1]$ is defined by

$$\varphi(x) := \mathbb{P}[g(x, \xi) \leq 0], \quad \text{for all } x \in \mathbb{R}^n. \quad (1)$$

Such functions are used to model uncertainty as probability constraints $\varphi(x) \geq p$, with a user-defined safety level $p \in (0, 1)$. Probability constrained optimization problems appear in fields such as energy, telecommunications, network expansion, mineral blending, chemical engineering; see, e.g., [1, 9, 15, 22, 25]. For a comprehensive overview on the theory and the applications of probabilistic constraints, we refer to [17, 18, 20].

The numerical treatment of the probabilistic constraints by optimization algorithms opens several questions, as convexity of the level set $\{x \in \mathbb{R}^n : \varphi(x) \geq p\}$ and smoothness of φ . Both (generalized) convexity (see e.g. [11, 12, 17, 18, 23]) and smoothness of φ have been the topic of several investigations. Regarding differentiability, the important results summarized in [21] prove that φ is continuously differentiable under the assumptions that $\{z \in \mathbb{R}^m : g(x, z) \leq 0\}$ is compact in a neighbourhood of x . Weaker conditions are shown to be sufficient in presence of special structure, for example separability of $g(x, \xi) = \xi - h(x)$, see [10]. For the general case of a nonlinear function g , the compactness assumption can also be replaced by growth conditions on the gradient $\nabla_x g$, as recently developed by [24] in presence of Gaussian multi-variate random vectors. To our knowledge, second-order differentiability has not been investigated yet – except for a specific case, given in the appendix of [25].

W. van Ackooij
EDF R&D. OSIRIS 1, avenue du Général de Gaulle, F-92141 Clamart Cedex France
E-mail: wim.van-ackooij@edf.fr

J. Malick
CNRS, LJK, F-38000 Grenoble, France
E-mail: jerome.malick@inria.fr

In this work, we study second-order differentiability of the probabilistic functions in the general framework of [24]. We prove that φ is twice continuously differentiable around a trial point \bar{x} , under growth conditions on $\nabla^2 g(x)$ the Hessian of g at x . We derive a general expression of $\nabla^2 \varphi(x)$ the Hessian of φ at x , as an integral with respect to the uniform distribution on the sphere. By Deák's sampling scheme of the sphere [4, 5], we can then approximate $\nabla^2 \varphi(x)$ as well as $\varphi(x)$ and $\nabla \varphi(x)$ simultaneously.

This expression of Hessians of probability functions could be useful to develop and analyze second-order methods for solving probabilistically constrained optimization problems. As shown for example in [26], first order optimization methods dealing with probabilistic constraints spend most time evaluating the gradients of the probability functions. Since we show here that Hessians are available at the same computational cost, second order methods (see e.g. [14]) could be considered, and this could speed up the overall resolution scheme. This may be particularly so for SQP methods that have only been investigated recently in this context, in [3].

Here is the outline of this short paper. Section 2 presents notation, background, and our second-order differentiability result. Then section 3 develops the technical lemmas leading to the proof of our result; this part relies heavily on the analysis of [24] regarding first-order differentiability. Finally section 4 presents examples and extensions.

2 (Second-order) differentiability under growth conditions

In this paper, we consider the general framework of [24] where g is a continuously differentiable function, convex with respect to the second argument, and ξ is a multivariate non-degenerate Gaussian random vector in \mathbb{R}^m . Without loss of generality¹, we can assume that $\xi \sim \mathcal{N}(0, R)$ is centered and with an arbitrary positive definite correlation matrix R . Note that the set $M(x) = \{z \in \mathbb{R}^m : g(x, z) \leq 0\}$ is convex so (Lebesgue) measurable and that ξ admits a density (with respect to the Lebesgue measure). Consequently the probability function φ is well-defined by (1).

Without additional assumptions though, φ is not differentiable in general (see a counterexample in [24, section 2]). The approach of [24] to get differentiability roughly requires to control the derivatives of g when z escapes to infinity. Specifically, the key assumption taken from [24] is the following first-order exponential growth condition, that we extend here to a second-order growth condition. This framework allows for many applications; see the examples in [24] and in the forthcoming section 4. Let us formalize here the two growth condition that will appear repeatedly in this paper.

Assumptions 1 (exponential growth conditions)

(i) *The mapping g satisfies the first order exponential growth condition at \bar{x} :*

$$\|\nabla_x g(x, z)\| \leq \delta \exp(\|z\|) \quad \text{for all } x \in U \text{ and all } z \text{ such that } \|z\| \geq C, \quad (2)$$

for a neighbourhood U of \bar{x} and constants $C, \delta > 0$. The norm $\|\cdot\|$ is the usual euclidian norm in \mathbb{R}^n .

(ii) *The mapping g satisfies the second order exponential growth condition at \bar{x} if it is twice continuously differentiable at \bar{x} and*

$$\|\nabla^2 g(x, z)\| \leq \delta \exp(\|z\|) \quad \text{for all } x \in U \text{ and all } z \text{ such that } \|z\| \geq C, \quad (3)$$

for a neighbourhood U of \bar{x} and constants $C, \delta > 0$. Here the norm of the Hessian refers to the operator norm induced by the usual euclidian norm of \mathbb{R}^n .

¹ Since we do not assume complex structure on g , the general case reduces to the normal centered case, as follows. Let $\tilde{\xi} \sim \mathcal{N}(\mu, \Sigma)$ be a general multi-variate Gaussian random vector with mean μ and covariance matrix Σ . We define $\tilde{g}(x, z) = g(x, Dz + \mu)$ and observe that $\xi = D^{-1}(\tilde{\xi} - \mu) \sim \mathcal{N}(0, R)$, where D is the diagonal matrix with $D_{ii} = \Sigma_{ii}^{1/2}$. We see that that $\mathbb{P}[g(x, \xi) \leq 0] = \mathbb{P}[\tilde{g}(x, \tilde{\xi}) \leq 0]$ and that neither convexity in the second argument of g , nor its continuous differentiability properties are perturbed by such a transformation.

Throughout this paper, we consider the Cholesky decomposition of R , i.e., $R = LL^\top$. We consider now the spherical-radial decomposition of ξ (e.g., [8, 16])

$$\xi = \eta L \zeta$$

with two independent random vectors: ζ having a uniform distribution over the Euclidian unit sphere \mathbb{S}^{m-1} in \mathbb{R}^m and η having a chi-distribution with m degrees of freedom. We denote by μ_ζ the uniform measure over the unit sphere \mathbb{S}^{m-1} and by μ_η the measure related to the chi-distribution with m degrees of freedom. With the factor $\kappa := 2^{1-\frac{m}{2}}/\Gamma(\frac{m}{2}) > 0$, the density associated to μ_η and its first derivative are

$$\chi(y) = \kappa y^{m-1} e^{-y^2/2} \quad \text{and} \quad \chi'(y) = \kappa((m-1) - y^2) y^{m-2} e^{-y^2/2}, \quad (4)$$

and the associated cumulative function is denoted F_χ . The mapping φ of (1) admits the following description as an integral

$$\varphi(x) = \int_{\{r \geq 0, v \in \mathbb{S}^{m-1} : g(x, rLv) \leq 0\}} \chi(r) dr d\mu_\zeta(v).$$

Given this description, the importance of studying the rays $\{r \geq 0, v \in \mathbb{S}^{m-1} : g(x, rLv) \leq 0\}$ becomes apparent. One can visualize a given $v \in \mathbb{S}^{m-1}$ as a direction along which we will investigate the length of a beam cutting the set $M(x)$. Since the set $M(x)$ is convex, such a beam can intersect the boundary of $M(x)$ at most twice. Under the additional assumption that $0 \in \text{int } M(x)$, such a beam intersects the boundary at most once. We will set aside these directions, called *finite* or *relevant* directions by introducing the following set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{S}^{m-1}$,

$$F(x) = \{v \in \mathbb{S}^{m-1} : \exists r > 0, g(x, rLv) = 0\}. \quad (5)$$

For an appropriate couple $x \in \mathbb{R}^n$ and $v \in F(x)$, the equation $g(x, rLv) = 0$ with respect to r can be solved locally around (\bar{x}, v) . This is exploited in the following nice alternative representation of φ , obtained as a consequence of Lemmas 3.1(4) and 3.3(1) of [24].

Theorem 1 (Integral representation) *In the above-described context, consider a point \bar{x} such that $g(\bar{x}, 0) < 0$. Then φ can be expressed in a neighbourhood of \bar{x} as*

$$\varphi(x) = \int_{v \in F(x)} F_\chi(\rho^{x,v}(x, v)) d\mu_\zeta(v) + \int_{v \notin F(x)} d\mu_\zeta(v). \quad (6)$$

where F is the set-valued function defined in (5), and $\rho^{x,v}$ is the real-valued function defined implicitly around (x, v) by the equation $g(x, rLv) = 0$ with respect to r .

We notice that the condition $g(\bar{x}, 0) < 0$ is not very restrictive as it holds in particular if $\varphi(\bar{x}) \geq \frac{1}{2}$ under a Slater assumption on g ; see [24, Proposition 3.11]. From the integral expression of this theorem, we see that differentiating φ involves differentiating under the integral a function defined implicitly. The main result of [24] gives general conditions that allow this differentiation.

Theorem 2 (First-order differentiability) *Let the assumption of Theorem 1 hold. Assume furthermore that g satisfies the first order exponential growth condition at \bar{x} (Assumption 1(i)). Then, φ is continuously differentiable on U , and for all $x \in U$*

$$\nabla \varphi(x) = - \int_{\substack{v \in F(x) \\ w = \rho^{x,v}(x, v)Lv}} \chi(\rho^{x,v}(x, v)) \nabla_x g(x, w) \alpha(x, v, w) d\mu_\zeta(v).$$

where α is the real-valued function defined by $\alpha(x, v, w) := 1/\langle \nabla_x g(x, w), Lv \rangle$.

We mention that, by [24, Theorem 3.14], the growth condition of this theorem can be replaced by the condition that the set $\{z : g(\bar{x}, z) \leq 0\}$ is bounded (and in this case, $\rho^{x,v}$ is in fact independent from (x, v) and $F(\bar{x}) = \mathbb{S}^{m-1}$ by [24, Lemma 3.12]).

In this short paper, we establish that a similar differentiability result holds at the second-order; more precisely here is our main second-order differentiability result.

Theorem 3 (Second-order differentiability) *Let the assumption of Theorem 1 hold. Assume furthermore that g satisfies the exponential growth conditions at \bar{x} (Assumption 1 (i) and (ii)). Then φ is twice continuously differentiable on a neighbourhood U of \bar{x} and for all $x \in U$*

$$\begin{aligned} \nabla^2 \varphi(x) = & \int_{\substack{v \in F(x) \\ w = \rho^{x,v}(x, v)Lv}} \left[\chi'(\rho^{x,v}(x, v)) \nabla_x g(x, w) \nabla_x g(x, w)^\top \right. \\ & + \chi(\rho^{x,v}(x, v)) \nabla_x g(x, w) (\nabla_{xz} g(x, w) Lv)^\top \\ & - \chi(\rho^{x,v}(x, v)) \alpha(x, v, w) v^\top L^\top \nabla_{zz} g(x, w) Lv \nabla_x g(x, w) \nabla_x g(x, w)^\top \\ & - \chi(\rho^{x,v}(x, v)) \alpha(x, v, w)^{-1} \nabla_{xx} g(x, w) \\ & \left. + \chi(\rho^{x,v}(x, v)) v^\top L^\top \nabla_{zx} g(x, w) \nabla_x g(x, w)^\top \right] \alpha(x, v, w)^2 d\mu_\zeta(v). \end{aligned} \quad (7)$$

Evaluating the Hessian $\nabla^2 \varphi(x)$ with the above integral expression can be done while evaluating $\varphi(x)$ and $\nabla \varphi(x)$ by Deàk's sampling method [5]. For each sampled point $v \in \mathbb{S}^{m-1}$, we try to solve the equation $g(x, rLv) = 0$ explicitly or approximately (by a Newton-Raphson algorithm for instance). If there is a solution $r > 0$ (i.e., $v \in F(x)$) then we evaluate the integrand; else (i.e., $v \notin F(x)$) the term does not contribute to the approximated integral. This procedure does not add significant computational cost on top of evaluating $\varphi(x)$ and $\nabla \varphi(x)$. Note that this is quite different from [25, Lemma 2], where the given expression of the Hessian (of a simple probability function) has a computing cost $O(m)$ times higher than the computing cost of the gradient.

3 Proof of second-order differentiability

This section provides several auxiliary developments that lead us to the proof of Theorem 3. We consider a given fixed, but arbitrary, point $\bar{x} \in \mathbb{R}^n$ such that $g(\bar{x}, 0) < 0$. The bulk of the work consists in studying the second-order differentiability of the function $e : \mathbb{R}^n \times \mathbb{S}^{m-1} \rightarrow [0, 1]$ defined by for all $x \in \mathbb{R}^n$ and $v \in \mathbb{S}^{m-1}$

$$e(x, v) := \mu_\eta(\{r \geq 0 : g(x, rLv) \leq 0\}) = \int_{\{r \geq 0 : g(x, rLv) \leq 0\}} \chi(r) dr \quad (8)$$

We start by giving more details on the implicit ray function $\rho^{\bar{x}, v}$ and the expressions of its first and second order derivatives (when $v \in F(\bar{x})$).

Lemma 1 (Implicit ray function) *If $\bar{v} \in F(\bar{x})$ (where F is defined as in (5)), then there exist neighbourhoods U of \bar{x} and V of \bar{v} as well as a twice continuously differentiable function $\rho^{\bar{x}, \bar{v}} : U \times V \rightarrow \mathbb{R}_+$ satisfying the equivalence*

$$g(x, rLv) = 0 \iff r = \rho^{\bar{x}, \bar{v}}(x, v) \quad \text{for all } (x, v, r) \in U \times V \times \mathbb{R}_+.$$

Moreover for all $(x, v) \in U \times V$ and for $w = \rho^{\bar{x}, \bar{v}}(x, v)Lv$

$$\nabla_x \rho^{\bar{x}, \bar{v}}(x, v) = -\alpha(x, v, w) \nabla_x g(x, w) \quad (9)$$

$$\begin{aligned} \nabla_{xx} \rho^{\bar{x}, \bar{v}}(x, v) = & \alpha(x, v, w)^2 \nabla_x g(x, w) (\nabla_{xz} g(x, w) Lv)^\top \\ & - \alpha(x, v, w)^3 v^\top L^\top \nabla_{zz} g(x, w) Lv \nabla_x g(x, w) \nabla_x g(x, w)^\top \\ & - \alpha(x, v, w) \nabla_{xx} g(x, w) + \alpha(x, v, w)^2 (v^\top L^\top \nabla_{zx} g(x, w) \nabla_x g(x, w)^\top). \end{aligned} \quad (10)$$

Proof The existence and smoothness of $\rho^{\bar{x}, \bar{v}}$ comes from applying the implicit function theorem. The details are in the proof of [24, Lemma 3.2], which also provides the expression of the gradient. We focus here on the calculation of the second order derivative. To this end let us consider $h(x, v) = \alpha(x, v, w)^{-1} = \langle \nabla_z g(x, \rho^{\bar{x}, \bar{v}}(x, v)Lv), Lv \rangle$. Then for any $k = 1, \dots, n$ we have the following derivative formula:

$$\begin{aligned} \frac{\partial h}{\partial x_k}(x, v) &= \frac{\partial}{\partial x_k} \left(\sum_{j=1}^m \frac{\partial g}{\partial z_j}(x, \rho^{\bar{x}, \bar{v}}(x, v)Lv)(Lv)_j \right) = \sum_{j=1}^m \frac{\partial^2 g}{\partial x_k \partial z_j}(x, \rho^{\bar{x}, \bar{v}}(x, v)Lv)(Lv)_j \\ &\quad + \sum_{j=1}^m (Lv)_j \sum_{\ell=1}^m \frac{\partial^2 g}{\partial z_\ell \partial z_j}(x, \rho^{\bar{x}, \bar{v}}(x, v)Lv)(Lv)_\ell \frac{\partial \rho^{\bar{x}, \bar{v}}}{\partial x_k}(x, v) \\ &= (\nabla_{xz} g(x, w)Lv)_k - \alpha(x, v, w) v^\top L^\top \nabla_{zz} g(x, w) Lv \frac{\partial g}{\partial x_k}(x, w). \end{aligned}$$

Consequently for any $k, i = 1, \dots, n$ we get:

$$\begin{aligned} \frac{\partial^2 \rho^{\bar{x}, \bar{v}}}{\partial x_i \partial x_k}(x, v) &= h(x, v)^{-2} \frac{\partial h}{\partial x_k}(x, v) \frac{\partial g}{\partial x_i}(x, \rho^{\bar{x}, \bar{v}}(x, v)Lv) - h(x, v)^{-1} \frac{\partial^2 g}{\partial x_i \partial x_k}(x, \rho^{\bar{x}, \bar{v}}(x, v)Lv) \\ &\quad + h(x, v)^{-1} \sum_{j=1}^m \frac{\partial^2 g}{\partial z_j \partial x_i}(x, \rho^{\bar{x}, \bar{v}}(x, v)Lv)(Lv)_j \frac{\partial \rho^{\bar{x}, \bar{v}}}{\partial x_k}(x, v), \end{aligned}$$

which by substitution gives

$$\begin{aligned} \frac{\partial^2 \rho^{\bar{x}, \bar{v}}}{\partial x_i \partial x_k}(x, v) &= h(x, v)^{-2} \frac{\partial g}{\partial x_i}(x, \rho^{\bar{x}, \bar{v}}(x, v)Lv) (\nabla_{xz} g(x, \rho^{\bar{x}, \bar{v}}(x, v)Lv)Lv)_k \\ &\quad - h(x, v)^{-3} v^\top L^\top \nabla_{zz} g(x, \rho^{\bar{x}, \bar{v}}(x, v)Lv) Lv \frac{\partial g}{\partial x_i}(x, \rho^{\bar{x}, \bar{v}}(x, v)Lv) \frac{\partial g}{\partial x_k}(x, \rho^{\bar{x}, \bar{v}}(x, v)Lv) \\ &\quad - h(x, v)^{-1} \frac{\partial^2 g}{\partial x_i \partial x_k}(x, \rho^{\bar{x}, \bar{v}}(x, v)Lv) \\ &\quad + h(x, v)^{-2} (v^\top L^\top \nabla_{zx} g(x, \rho^{\bar{x}, \bar{v}}(x, v)Lv))_i \frac{\partial g}{\partial x_k}(x, \rho^{\bar{x}, \bar{v}}(x, v)Lv). \end{aligned}$$

This can be written as (10). □

The second-order differentiability of e when $\bar{v} \in F(\bar{x})$ follows immediately from the previous lemma together with Corollary 3.5 of [24].

Lemma 2 (In relevant directions) *If $\bar{v} \in F(\bar{x})$, then there exist neighbourhoods U of \bar{x} and V of \bar{v} such that, for $(x, v) \in U \times V$, the function e defined in (8) satisfies $e(x, v) = F_\chi(\rho^{\bar{x}, \bar{v}}(x, v))$ as well as*

$$\frac{\partial e}{\partial x_k}(x, v) = -\alpha(x, v, w) \chi(\rho^{\bar{x}, \bar{v}}(x, v)) \frac{\partial g}{\partial x_k}(x, w)$$

and is moreover twice continuously differentiable with partial derivative:

$$\frac{\partial^2 e}{\partial x_k \partial x_i}(x, v) = \chi'(\rho^{\bar{x}, \bar{v}}(x, v)) \alpha(x, v, w)^2 \frac{\partial g}{\partial x_k}(x, w) \frac{\partial g}{\partial x_i}(x, w) + \chi(\rho^{\bar{x}, \bar{v}}(x, v)) \frac{\partial^2 \rho^{\bar{x}, \bar{v}}}{\partial x_k \partial x_i}(x, v),$$

where $k, i \in \{1, \dots, n\}$ are arbitrary and $w = \rho^{\bar{x}, \bar{v}}(x, v)Lv$.

We now show that Assumptions 1 yield that the exponentials of the χ -distribution make the Hessian of e vanish when getting close to irrelevant directions ($\bar{v} \notin F(\bar{x})$).

Lemma 3 (Toward irrelevant directions) *Let $\bar{v} \notin F(\bar{x})$ and consider a sequence $(x_k, v_k) \rightarrow (\bar{x}, \bar{v})$ with $v_k \in F(x_k)$. If g satisfies Assumption 1 (i) and (ii) at \bar{x} , then*

$$\lim_{k \rightarrow \infty} \nabla_{xx} e(x_k, v_k) = 0.$$

Proof By [24, Lemma 3.3], we have first that $\rho_k := \rho^{x_k, v_k}(x_k, v_k) \rightarrow \infty$ when $k \rightarrow \infty$, since $x_k \rightarrow \bar{x}$ and $v_k \in F(x_k)$. Recall also that, by [24, Lemma 3.1(3)], there exists $\delta_1 > 0$ such that, on an appropriately small neighbourhood of \bar{x} ,

$$\langle \nabla_z g(x_k, \rho_k L v_k), L v_k \rangle \geq \delta_1 [\rho^{x_k, v_k}(x_k, v_k)]^{-1} = \delta_1 \rho_k^{-1} > 0.$$

Now note that, by Lemma 2, we have the upper bound:

$$\|\nabla_{xx} e(x_k, v_k)\| \leq \left| \frac{\chi'(\rho_k)}{\langle \nabla_z g(x_k, \rho_k L v_k), L v_k \rangle^2} \right| \|\nabla_x g(x_k, \rho_k L v_k)\|^2 + \chi(\rho_k) \|\nabla_{xx} \rho^{x_k, v_k}(x_k, v_k)\|. \quad (11)$$

By using the growth condition for k sufficiently large and the expressions of (4), we can bound the first term (which we call A_k for short) by

$$\begin{aligned} A_k &\leq \delta^2 \delta_1^{-2} \rho_k^2 \kappa \{ (m-1) - \rho_k^2 \} [\rho_k]^{m-2} e^{-[\rho_k]^2/2} e^{2\|L\|\rho_k} \\ &= \delta \kappa \delta_1^{-2} \kappa \{ (m-1) - \rho_k^2 \} [\rho_k]^m e^{2\|L\|\rho_k} e^{-[\rho_k]^2/2}, \end{aligned}$$

which allows us to show that $\lim_{k \rightarrow \infty} A_k = 0$ since $\rho_k := \rho^{x_k, v_k}(x_k, v_k) \rightarrow \infty$ and $y^\alpha e^y e^{-y^2/2} \rightarrow 0$ for $y \rightarrow \infty$, where $\alpha > 0$ is arbitrary. We bound the second term in (11) by using the second order growth condition:

$$\begin{aligned} &\chi(\rho^{x_k, v_k}(x_k, v_k)) \|\nabla_{xx} \rho^{x_k, v_k}(x_k, v_k)\| \\ &\leq \kappa \delta_1^{-1} \left[\delta_1^{-1} \rho_k^2 \|L\| e^{2\|L\|\rho_k} [\rho_k]^{m-1} e^{-[\rho_k]^2/2} + \delta_1^{-2} \rho_k^3 \|L\|^2 e^{3\|L\|\rho_k} [\rho_k]^{m-1} e^{-[\rho_k]^2/2} \right. \\ &\quad \left. + \rho_k e^{\|L\|\rho_k} [\rho_k]^{m-1} e^{-[\rho_k]^2/2} + \delta_1^{-1} \rho_k^2 \|L\| e^{2\|L\|\rho_k} [\rho_k]^{m-1} e^{-[\rho_k]^2/2} \right]. \end{aligned}$$

The right-hand side can be seen to converge to zero following the above evoked arguments. \square

We are now in position to establish the twice differentiability of e , which is the main technical lemma in our way to Theorem 3.

Lemma 4 (Second derivative of the function e) *If g satisfies the first and second order exponential growth conditions at \bar{x} , then there is a neighbourhood U of \bar{x} such that the mapping e is twice continuously differentiable with respect to x on U . For $x \in U$ and $v \notin F(x)$, the gradient and Hessian of e are null, and for $v \in F(x)$, the gradient is given in Lemma 2 and the Hessian by*

$$\begin{aligned} \nabla_{xx} e(x, v) &= \alpha(x, v, w)^2 \left(\chi'(\rho^{x, v}(x, v)) \nabla_x g(x, w) \nabla_x g(x, w)^\top \right. \\ &\quad + \chi(\rho^{x, v}(x, v)) \nabla_x g(x, w) (\nabla_{xz} g(x, w) L v)^\top \\ &\quad - \chi(\rho^{x, v}(x, v)) \alpha(x, v, w) v^\top L^\top \nabla_{zz} g(x, w) L v \nabla_x g(x, w) \nabla_x g(x, w)^\top \\ &\quad - \chi(\rho^{x, v}(x, v)) / \alpha(x, v, w) \nabla_{xx} g(x, w) \\ &\quad \left. + \chi(\rho^{x, v}(x, v)) v^\top L^\top \nabla_{zx} g(x, w) \nabla_x g(x, w)^\top \right), \end{aligned}$$

where $w = \rho^{x, v}(x, v) L v$. We have moreover that $\nabla_{xx} e$ is continuous with respect to the couple (x, v) .

Proof Fix $i, \ell \in \{1, \dots, n\}$. For any $v \in F(x)$, the provided formula and differentiability statement follows from Lemma 2. We focus here on the case $v \notin F(x)$. We will show, by contradiction, that,

$$\lim_{t \uparrow 0} \frac{\frac{\partial e}{\partial x_\ell}(x + t u_i, v) - \frac{\partial e}{\partial x_\ell}(x, v)}{t} = 0, \quad (12)$$

where u_i is the i -th canonical unit vector in \mathbb{R}^n . In exactly the same way, one can also show that the limit for $t \downarrow 0$ equals zero too. Altogether, this will prove that e is twice differentiable at x and that $\frac{\partial^2 e}{\partial x_i \partial x_\ell} e(x, v) = 0$.

Notice first that, by [24, Corollary 3.8], we have $\frac{\partial e}{\partial x_\ell}(x, v) = 0$. Negating (12) implies the existence of some $\varepsilon > 0$ and a subsequence $t_k \uparrow 0$ such that

$$\left| \frac{\frac{\partial e}{\partial x_\ell}(x + t_k u_i, v)}{t_k} \right| \geq \varepsilon. \quad (13)$$

This implies in particular that $\frac{\partial e}{\partial x_\ell}(x + t_k u_i, v) \neq 0$ and then $v \in F(x + t_k u_i)$, again by [24, Corollary 3.8]. We also have that $g(x + t_k u_i, 0) < 0$ for all large k . Now, for an arbitrary k , we define (recall that $t_k < 0$)

$$T := \inf \left\{ \tau \in [t_k, 0] : \frac{\partial e}{\partial x_\ell}(x + \tau u_i, v) = 0 \right\} \quad (\leq 0).$$

Due to $\frac{\partial e}{\partial x_\ell}(x, v) = 0$ we have that $T \leq 0$. On the other hand, $\frac{\partial e}{\partial x_\ell}(x + t_k u_i, v) \neq 0$ and the continuity of $\nabla_x e$ established in [24, Corollary 3.9] provides $T > t_k$. We infer that $\frac{\partial e}{\partial x_\ell}(x + \tau u_i, v) \neq 0$ for all $\tau \in [t_k, T)$ and, hence,

$$v \in F(x + \tau u_i) \quad \forall \tau \in [t_k, T). \quad (14)$$

But then, the function

$$\beta(\tau) := \frac{\partial e}{\partial x_\ell}(x + \tau u_i, v)$$

is differentiable for all $\tau \in (t_k, T)$ by virtue of Lemma 2 and its derivative is given by

$$\beta'(\tau) = \sum_{j=1}^n \frac{\partial^2 e}{\partial x_j \partial x_\ell}(x + \tau u_i, v) u_i(j) = \frac{\partial^2 e}{\partial x_i \partial x_\ell}(x + \tau u_i, v).$$

Therefore, the mean value theorem guarantees the existence of some $\tau_k^* \in (t_k, T)$ such that

$$\beta'(\tau_k^*) = \frac{\beta(T) - \beta(t_k)}{T - t_k}.$$

By continuity of $\nabla_x e$ of [24, Corollary 3.9], we have that $\beta(T) = \frac{\partial e}{\partial x_\ell}(x + T u_i, v) = 0$, whence, by $t_k < T \leq 0$,

$$|\beta'(\tau_k^*)| = \left| \frac{-\frac{\partial e}{\partial x_\ell}(x + t_k u_i, v)}{T - t_k} \right| \geq \left| \frac{\frac{\partial e}{\partial x_\ell}(x + t_k u_i, v)}{t_k} \right| \geq \varepsilon,$$

where the last inequality is (13). Now, since k was arbitrarily fixed, we have constructed a sequence τ_k^* such that $t_k < \tau_k^* \leq 0$ such that

$$\left| \frac{\partial^2 e}{\partial x_i \partial x_\ell}(x + \tau_k^* u_i, v) \right| \geq \varepsilon \quad \forall k. \quad (15)$$

Since $t_k \uparrow 0$, we also have that $\tau_k^* \uparrow 0$. Moreover, $v \in F(x + \tau_k^* u_i)$ by (14). Due to our assumption that g satisfies the growth conditions at x , Lemma 3 yields that $\lim_{k \rightarrow \infty} \nabla_{xx} e(x_k, v) = 0$ which contradicts (15) and thus concludes about twice differentiability. Finally, the fact that the second order partial derivative $\nabla_{xx} e$ is also continuous at (x, v) for any $v \in \mathbb{S}^{m-1}$ can be shown as follows. For a given x and $v \in F(x)$, the provided formula of the Hessian holds true locally and continuity is evident. In contrast, if $v \notin F(x)$, assuming that $\nabla_{xx} e(x_k, v_k)$ does not tend to zero along a sequence $(x_k, v_k) \rightarrow (x, v)$, means in particular that $\|\nabla_{xx} e(x_k, v_k)\| \geq \varepsilon$ for a given $\varepsilon > 0$. Consequently by the arguments above, $v_k \in F(x_k)$. But then this leads to a contradiction with Lemma 3. \square

The proof of Theorem 3 about second-order differentiability of φ now follows from the previous result after justifying the derivation under the integral.

Proof (of Theorem 3) Theorem 1 gives the integral representation (6) of φ around \bar{x} . Theorem 2 allows us to obtain an integral representation for $\nabla\varphi(x)$ involving e . Indeed for a given $k = 1, \dots, n$, we have

$$\frac{\partial\varphi}{\partial x_k}(x) = \int_{v \in \mathbb{S}^{m-1}} \frac{\partial e}{\partial x_k}(x, v) d\mu_\zeta(v), \quad (16)$$

as a consequence of Theorem 2 and expressions of derivatives of e of Lemmas 2 and 4. Invoking again Lemma 4, the second order partial derivative $\nabla_{xx}e$ of the function e exists and is continuous on $U \times \mathbb{S}^{m-1}$. The continuity of $\nabla_{xx}e$ on $U \times \mathbb{S}^{m-1}$ and compactness of \mathbb{S}^{m-1} guarantee that the function $\gamma : U \rightarrow \mathbb{R}$

$$\gamma(x) := \max_{v \in \mathbb{S}^{m-1}} \|\nabla_{xx}e(x, v)\|$$

is well-defined and continuous. Restricting U if necessary, we may assume that $\gamma(x) \leq 2\gamma(\bar{x})$ for all $x \in U$. This together with $\mu_\zeta(\mathbb{S}^{m-1}) = 1$ for the law μ_ζ of the uniform distribution on \mathbb{S}^{m-1} we infer that the constant $2\gamma(\bar{x})$ is an integrable function on \mathbb{S}^{m-1} dominating $\|\nabla_{xx}e(x, v)\|$ on \mathbb{S}^{m-1} for all $x \in U$. Now, Lebesgue's dominated convergence theorem [13, Theorem 12.35, Corollary 12.36] allows to differentiate (16) under the integral sign and provides continuity with respect to x . We have established

$$\nabla^2\varphi(x) = \int_{v \in \mathbb{S}^{m-1}} \nabla_{xx}e(x, v) d\mu_\zeta$$

which gives the desired expression by Lemma 4. \square

4 Examples and extensions

We illustrate now our second-order differentiability result on some special examples: for chi-squared random vectors in section 4.1 and for lognormal random vectors in section 4.2. Finally we show in section 4.3 that results similar to Theorem 3 also hold with Student random vectors.

4.1 Chi-squared example

We consider here a linear probabilistic function

$$\varphi(x) = \mathbb{P}(\langle \eta, x \rangle \leq b) \quad (17)$$

with $b > 0$ and a random vector $\eta \in \mathbb{R}^n$ whose components η_i are independent and have a χ^2 -distribution with m_i degrees of freedom. For each $i = 1, \dots, n$, we can write $\eta_i = \sum_{k=1}^{m_i} \xi_{i,k}^2$, where m_i is the degree of freedom, and $\xi_{i,k} \sim \mathcal{N}(0, 1)$ (for $k = 1, \dots, m_i$) are independent random variables. We consider the Gaussian random vector of $m = m_1 + \dots + m_n$ dimensions

$$\xi := (\xi_{1,1}, \dots, \xi_{1,m_1}, \dots, \xi_{n,1}, \dots, \xi_{n,m_n}) \sim \mathcal{N}(0, I).$$

We also introduce the function $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ defined by its components $f_i(z) = \sum_{k=1}^{m_i} z_{i,k}^2$ (for z partitioned as ξ above), and the function $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ defined by $g(x, z) = \langle x, f(z) \rangle - b$. Thus, we are in the situation of this paper as the probability function (17) writes

$$\varphi(x) = \mathbb{P}(\langle \eta, x \rangle \leq b) = \mathbb{P}(\langle f(\xi), x \rangle \leq b) = \mathbb{P}(g(x, \xi) \leq 0).$$

We observe that g is C^2 , satisfies the first and second order growth conditions and is convex with respect to z as soon as $x_i \geq 0$ for all $i = 1, \dots, n$. We apply Theorem 3 at \bar{x} (such that $\bar{x}_i > 0$ for all $i = 1, \dots, n$) to establish that φ of (17) is C^2 around \bar{x} , and we can specify the integral expression of the Hessian for this case as follows. We observe that the set $\{z \in \mathbb{R}^m : g(\bar{x}, z) \leq 0\}$ is bounded, so that $F(\bar{x}) = \mathbb{S}^{m-1}$ by [24, Lemma 3.12]. We also dispose of an explicit ray function (for all (x, v)) as the equation $\langle f(rv), x \rangle = b$ in r admits the solution

$$\rho(x, v) = \sqrt{b / \langle f(v), x \rangle}.$$

We readily compute the terms that appeared in the expression of Theorem 3, including

$$\alpha(x, v, \rho(x, v)v) = 1/(2\sqrt{b\langle f(v), x \rangle}) \quad \text{and} \quad v^\top \nabla_{zz} g(x, \rho(x, v)v) = 2\langle f(v), x \rangle.$$

Substituting these expressions in (7) and simplifying yields:

$$\nabla^2 \varphi(x) = \int_{v \in \mathbb{S}^{m-1}} \left[b \chi' \left(\sqrt{b\langle f(v), x \rangle} \right) + 3\sqrt{b\langle f(v), x \rangle} \chi \left(\sqrt{b\langle f(v), x \rangle} \right) \right] \frac{f(v)^\top f(v)}{4\langle f(v), x \rangle^3} d\mu_\zeta(v).$$

4.2 Multivariate log-normal distribution

We consider here the probabilistic function

$$\varphi(x) = \mathbb{P}(\langle \eta, x \rangle \leq h(x)) \quad (18)$$

with a twice continuously differentiable $h : \mathbb{R}^m \rightarrow \mathbb{R}$ and a random vector η with a multivariate lognormal distribution, i.e., the component-wise logarithm $\log \eta$ has a multivariate Gaussian distribution). Without loss of generality we assume that $\xi = \log \eta \sim \mathcal{N}(0, R)$ for some correlation matrix $R = LL^\top$. Thus we remark that the probabilistic function of (18) can be written as $\varphi(x) = \mathbb{P}[g(x, \xi) \leq 0]$ with the function $g(x, z) = \langle x, e^z \rangle - h(x)$, which is C^2 , satisfies the first and second order growth conditions, and is convex with respect to z as soon as $x_i \geq 0$ for all $i = 1, \dots, m$.

Let \bar{x} be such that $\bar{x}_i > 0$ for $i = 1, \dots, m$ and $h(\bar{x}) > \sum_{i=1}^m \bar{x}_i$. This implies that $g(\bar{x}, 0) < 0$ and we can therefore apply Theorem 3, to get that φ of (18) is C^2 around \bar{x} (more precisely, in a neighbourhood U of \bar{x} such that $g(x, 0) < 0$ and $x_i > 0$ for $i = 1, \dots, m$). The expression of α and F instantiates as

$$\alpha(x, v, w) = \left(\sum_{i=1}^m x_i e^{\rho^{x,v}(x,v)(Lv)_i} (Lv)_i \right)^{-1} \quad \text{and} \quad F(x) = \{v \in \mathbb{S}^{m-1} : \exists i \text{ such that } (Lv)_i > 0\}. \quad (19)$$

The integral expression of the Hessian then simplifies to

$$\begin{aligned} \nabla^2 \varphi(x) = & \int_{v \in F(x)} \left(\chi'(\rho^{x,v}(x,v)) \left[e^{\rho^{x,v}(x,v)Lv} - \nabla h(x) \right] \left[e^{\rho^{x,v}(x,v)Lv} - \nabla h(x) \right]^\top \right. \\ & + \chi(\rho^{x,v}(x,v)) \left[e^{\rho^{x,v}(x,v)Lv} - \nabla h(x) \right] \left[\text{diag}(e^{\rho^{x,v}(x,v)Lv})Lv \right]^\top \\ & - \frac{\chi(\rho^{x,v}(x,v))}{\alpha(x, v, \rho^{x,v}(x,v))} v^\top L^\top \text{diag}(x e^{\rho^{x,v}(x,v)Lv})Lv \left[e^{\rho^{x,v}(x,v)Lv} - \nabla h(x) \right] \left[e^{\rho^{x,v}(x,v)Lv} - \nabla h(x) \right]^\top \\ & + \chi(\rho^{x,v}(x,v)) \alpha(x, v, \rho^{x,v}(x,v)) \nabla^2 h(x) \\ & \left. + \chi(\rho^{x,v}(x,v)) \left[v^\top L^\top \text{diag}(e^{\rho^{x,v}(x,v)Lv}) \right] \left[e^{\rho^{x,v}(x,v)Lv} - \nabla h(x) \right]^\top \right) \alpha(x, v, w)^2 d\mu_\zeta(v). \end{aligned}$$

Here, e^z for a vector z has to be understood component-wise, $\text{diag}(y)$ denotes the diagonal matrix with vector y on the diagonal, and $x e^z$ has to be understood as the components wise product.

The expression of $F(x)$ in (19) is the main technical point to applying Theorem 3 to the example of this section. For sake of completeness, let us provide a quick proof of this expression: we show by double inclusion that, for all $x \in U$,

$$\mathbb{S}^{m-1} \setminus F(x) = \{v \in \mathbb{S}^{m-1} : Lv \leq 0\}. \quad (20)$$

First, let $x \in U$ and $v \in \mathbb{S}^{m-1}$ with $Lv \leq 0$; then for all $r > 0$

$$g(x, rLv) = \langle e^{rLv}, x \rangle - h(x) \leq \langle e^0, x \rangle - h(x) = g(x, 0) < 0,$$

whence $v \notin F(x)$. Conversely, let $x \in U$ and $v \notin F(x)$. Then, $\langle e^{rLv}, x \rangle < h(x)$ for all $r > 0$. Define $J := \{i : (Lv)_i > 0\}$ and observe that $\sum_{i \in J} x_i e^{r(Lv)_i}$ is bounded from above independently of r . If $J \neq \emptyset$, then this sum would tend to ∞ for $r \rightarrow \infty$ which is a contradiction. Consequently, $J = \emptyset$, proving $Lv \leq 0$ and, thus, the reverse inclusion of (20).

4.3 Second-order differentiability with Student's law

The rationale of the proof of section 3 can be adapted to a similar context under slight changes in conditions and analysis. In this section, we show for example that the second-order differentiability still holds when the random vector ξ in (1) follows a multi-variate Student distribution, under a slightly adapted growth condition on the Hessian of g .

Consider the probability function φ defined in (1) where $\xi \sim \mathcal{T}(0, R, \nu)$ has a standard multi-variate Student distribution (with correlation matrix $R = LL^\top$ and ν degrees of freedom). By [24, Theorem 4.4], φ can be expressed as

$$\varphi(x) = \int_{v \in \mathbb{S}^{m-1}} \tilde{e}(x, v) d\mu_\zeta(v) \quad \text{with} \quad \tilde{e}(x, v) = \begin{cases} F_{m, \nu}(m^{-1} [\rho^{x, v}(x, v)]^2) & \text{if } v \in F(x) \\ 1 & \text{else} \end{cases}, \quad (21)$$

where $F_{m, \nu}(t)$ is the Fisher-Snedecor distribution function with density:

$$f_{m, \nu}(t) = \begin{cases} \frac{\Gamma(m/2 + \nu/2)}{\Gamma(m/2)\Gamma(\nu/2)} m^{m/2} \nu^{\nu/2} t^{m/2-1} (mt + \nu)^{-(m+\nu)/2} & t \geq 0 \\ 0 & t < 0 \end{cases} \quad (22)$$

Theorem 4 *Let the mapping $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ be twice continuously differentiable and convex in the second argument. Assume moreover that g satisfies a first and second order polynomial growth condition at \bar{x} (i.e., the exponential term in (3) is replaced with $\|z\|^\ell$, and the one of (2) by $\|z\|^\varkappa$ and these coefficients are related as follows $\ell + 2\varkappa < \nu - 2$). Then φ of (21) is twice continuously differentiable on a neighbourhood U of \bar{x} and for all $x \in U$*

$$\begin{aligned} \nabla^2 \varphi(x) = & \int_{\substack{v \in F(x) \\ w = \rho^{x, v}(x, v)Lv}} \left[4 \frac{\rho^{x, v}(x, v)^2 f'_{m, \nu}(m^{-1} \rho^{x, v}(x, v)^2)}{m f_{m, \nu}(m^{-1} \rho^{x, v}(x, v)^2)} \nabla_x g(x, w) \nabla_x g(x, w)^\top \right. \\ & + 2 \nabla_x g(x, w) \nabla_x g(x, w)^\top + 2 \rho^{x, v}(x, v) \nabla_x g(x, w) (\nabla_{xz} g(x, w) Lv)^\top \\ & - 2 \rho^{x, v}(x, v) \alpha(x, v, w) v^\top L^\top \nabla_{zz} g(x, w) Lv \nabla_x g(x, w) \nabla_x g(x, w)^\top \\ & \left. - 2 \rho^{x, v}(x, v) / \alpha(x, v, w) \nabla_{xx} g(x, w) \right. \\ & \left. + 2 \rho^{x, v}(x, v) v^\top L^\top \nabla_{zx} g(x, w) \nabla_x g(x, w)^\top \right] f_{m, \nu} \left(\frac{\rho^{x, v}(x, v)^2}{m} \right) \frac{\alpha(x, v, w)^2}{m} d\mu_\zeta(v) \end{aligned}$$

Proof The continuous differentiability is established in [24, Theorem 4.7]. We focus on the proof of the second-order differentiability and, as remarked in [24], we only need to derive the equivalent of Lemma 3 for the Student setting. Following the arguments of Lemma 3, we have

$$\begin{aligned} \|\nabla_{xx} \tilde{e}(x_k, v_k)\| \leq & \left| \frac{4 \rho^{x, v}(x, v)^2 f'_{m, \nu}(m^{-1} [\rho^{x, v}(x, v)]^2)}{m^2 \langle \nabla_z g(x_k, \rho^{x_k, v_k}(x_k, v_k) Lv_k), Lv_k \rangle^2} \right| \|\nabla_x g(x_k, \rho^{x_k, v_k}(x_k, v_k) Lv_k)\|^2 \\ & + \left| \frac{2 f_{m, \nu}(m^{-1} [\rho^{x, v}(x, v)]^2)}{m \langle \nabla_z g(x_k, \rho^{x_k, v_k}(x_k, v_k) Lv_k), Lv_k \rangle^2} \right| \|\nabla_x g(x_k, \rho^{x_k, v_k}(x_k, v_k) Lv_k)\|^2 \\ & + \frac{2}{m} \rho^{x, v}(x, v) f_{m, \nu}(m^{-1} [\rho^{x, v}(x, v)]^2) \|\nabla_{xx} \rho^{x_k, v_k}(x_k, v_k)\|. \end{aligned}$$

Using the expression of $f'_{m, \nu}(t)$ (for $t \geq 0$)

$$f'_{m, \nu}(t) = \frac{\Gamma(m/2 + \nu/2)}{\Gamma(m/2)\Gamma(\nu/2)} m^{m/2} \nu^{\nu/2} t^{m/2-1} (mt + \nu)^{-(m+\nu)/2} \left(\left(\frac{m}{2} - 1 \right) t^{-1} - m \frac{m + \nu}{2} (mt + \nu)^{-1} \right)$$

and the polynomial growth condition, we can bound the first term (A_k) as:

$$\|A_k\| \leq \tilde{C}_1 \rho_k^{m+2\varkappa} (\rho_k^2 + \nu)^{-\frac{m+\nu}{2}} + \tilde{C}_2 \rho_k^{m+2\varkappa+2} (\rho_k^2 + \nu)^{-\frac{m+\nu+2}{2}},$$

for appropriate constants \tilde{C}_1, \tilde{C}_2 the last term tends to zero whenever $\varkappa < \frac{\nu}{2}$ and where ρ_k is short for $\rho_k := \rho^{x_k, v_k}(x_k, v_k)$. The second term (B_k) can be bounded by:

$$\|B_k\| \leq \tilde{C}_3 \rho^{x_k, v_k}(x_k, v_k)^{m+2\varkappa} (\rho^{x_k, v_k}(x_k, v_k)^2 + \nu)^{-\frac{m+\nu}{2}}$$

which again tends to zero when $k \rightarrow \infty$ as long as $\varkappa < \frac{\nu}{2}$. Finally the last term (C_k) can be bounded by using the second order polynomial growth condition as follows:

$$\|C_k\| \leq \tilde{C}_4 \rho_k^m (\rho_k^2 + \nu)^{-\frac{m+\nu}{2}} (2\rho_k^{1+\varkappa+\ell} + \rho_k^{2+2\varkappa+\ell} + \rho_k^\ell).$$

When $\ell + 2\varkappa < \nu - 2$, we also have $\ell < \nu$, $1 + \varkappa + \ell < \nu$ and $\varkappa < \frac{\nu}{2}$, so that this condition entails that also $\|C_k\| \rightarrow 0$, when $k \rightarrow \infty$. We have thus established the equivalent of Lemma 3: the mapping \tilde{e} has the same properties as e . Consequently the remaining proofs carry forth. \square

References

1. T. Arnold, R. Henrion, A. Möller, and S. Vigerske. A mixed-integer stochastic nonlinear optimization problem with joint probabilistic constraints. *Pacific Journal of Optimization*, 10:5–20, 2014.
2. M. Bertocchi, G. Consigli, and M.A.H. Dempster (Eds). *Stochastic Optimization Methods in Finance and Energy: New Financial Products and Energy Market Strategies*. International Series in Operations Research and Management Science. Springer, 2012.
3. I. Bremer, R. Henrion, and A. Möller. Probabilistic constraints via SQP solver: Application to a renewable energy management problem. *Computational Management Science*, 12:435–459, 2015.
4. I. Deák. Computing probabilities of rectangles in case of multinormal distribution. *Journal of Statistical Computation and Simulation*, 26(1-2):101–114, 1986.
5. I. Deák. Subroutines for computing normal probabilities of sets - computer experiences. *Annals of Operations Research*, 100:103–122, 2000.
6. J. Elstrodt. *Maß und Integrationstheorie*. Springer-Verlag, 7th edition, 2011.
7. C.A. Floudas and P.M. Pardalos (Eds). *Encyclopedia of Optimization*. Springer - Verlag, 2nd edition, 2009.
8. A. Genz and F. Bretz. *Computation of multivariate normal and t probabilities*. Number 195 in Lecture Notes in Statistics. Springer, Dordrecht, 2009.
9. R. Henrion and A. Möller. Optimization of a continuous distillation process under random inflow rate. *Computer & Mathematics with Applications*, 45:247–262, 2003.
10. R. Henrion and A. Möller. A gradient formula for linear chance constraints under Gaussian distribution. *Mathematics of Operations Research*, 37:475–488, 2012.
11. R. Henrion and C. Strugarek. Convexity of chance constraints with independent random variables. *Computational Optimization and Applications*, 41:263–276, 2008.
12. R. Henrion and C. Strugarek. *Convexity of Chance Constraints with Dependent Random Variables: the use of Copulae*. (Chapter 17 in [2]). Springer New York, 2011.
13. J. K. Hunter and B. Nachtergaele. *Applied Analysis*. World Scientific Publishing Company, 2001.
14. A. F. Izmailov and M. V. Solodov. *Newton-Type Methods for Optimization and Variational Problems*. Springer Series in Operations Research and Financial Engineering. Springer, 2014.
15. D.R. Morgan, J.W. Eheart, and A.J. Valocchi. Aquifer remediation design under uncertainty using a new chance constraint programming technique. *Water Resources Research*, 29:551–561, 1993.
16. A. Naor and D. Romik. Projecting the surface measure of the sphere of ℓ_p^n . *Ann. I.H. Poincaré*, 39(2):241–261, 2003.
17. A. Prékopa. *Stochastic Programming*. Kluwer, Dordrecht, 1995.
18. A. Prékopa. *Probabilistic programming*. In [19] (Chapter 5). Elsevier, Amsterdam, 2003.
19. A. Ruszczyński and A. Shapiro. *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*. Elsevier, Amsterdam, 2003.
20. A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming. Modeling and Theory*, volume 9 of *MPS-SIAM series on optimization*. SIAM and MPS, Philadelphia, 2009.
21. S. Uryas'ev. *Derivatives of probability and Integral functions: General Theory and Examples*. Appearing in [7]. Springer - Verlag, 2nd edition, 2009.
22. W. van Ackooij. Decomposition approaches for block-structured chance-constrained programs with application to hydro-thermal unit commitment. *Mathematical Methods of Operations Research*, 80(3):227–253, 2014.
23. W. van Ackooij. Eventual convexity of chance constrained feasible sets. *Optimization (A Journal of Math. Programming and Operations Research)*, 64(5):1263–1284, 2015.
24. W. van Ackooij and R. Henrion. Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distributions. *SIAM Journal on Optimization*, 24(4):1864–1889, 2014.
25. W. van Ackooij, R. Henrion, A. Möller, and R. Zorgati. Joint chance constrained programming for hydro reservoir management. *Optimization and Engineering*, 15:509–531, 2014.
26. W. van Ackooij and C. Sagastizábal. Constrained bundle methods for upper inexact oracles with application to joint chance constrained energy problems. *SIAM Journal on Optimization*, 24(2):733–765, 2014.